Particle Swarms in Statistical Physics

Andrei Băutu and Elena Băutu

1 “Mircea cel Batrân” Naval Academy, 2 “Ovidius” University
Romania

1. Introduction

Statistical physics is the area of physics which studies the properties of systems composed of many microscopic particles (like atoms and molecules). When combined, the interactions between these particles produce the macroscopic features of the systems. The systems are usually characterized by a very large number of variables and the limited possibilities for observing the properties of the components of the system. For these reasons, solving problems arisen in Statistical Physics with analytical approaches is usually ineffective and sometimes impossible. However, statistical approaches (such as Monte Carlo simulation) can provide acceptable approximations for solutions of these problems. Moreover, recent studies showed that nature inspired metaheuristics (like Genetic Algorithms, Evolutionary Strategies, Particle Swarm Optimization, etc) can also be used to simulate, analyse, and optimize such systems, providing fast and accurate results. Apart from physical implications, problems from Statistical Physics are also important in fields like biology, chemistry, mathematics, communications, economy, sociology, etc.

We will present two important problems from Statistical Physics and discuss how one can use Particle Swarm Optimization (PSO) to tackle them. First, we will discuss how the real-valued version of PSO can be used to minimize the energy of a system composed of repulsive point charges confined on a sphere. This is known as the Thomson problem and it is included in Stephen Smale's famous list of 18 unsolved mathematical problems to be solved in the 21st century. This problem also arises in biology, chemistry, communications, economy, etc.

Latter on, we will discuss how the binary version of PSO can be used to search ground states of Ising spin glasses. Spin glasses are materials that simultaneously present ferromagnetic and anti-ferromagnetic interactions among their atoms. A ground state of a spin glass is a configuration of the system in which this has the lowest energy possible. Besides its importance for Statistical Physics, this problem has applications in neural network theory, computer science, biology, etc.

2. The Basics of Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a metaheuristic inspired by the behaviour of social creatures, which interact between them in order to achieve a common goal. Such behaviour...
can be noticed in flocks of birds searching for food or schools of fish trying to avoid predators (Eberhart & Kennedy, 1995).

The philosophy of PSO is based on the evolutionary cultural model, which states that in social environments individuals have two learning sources: individual learning and cultural transmission (Boyd & Richerson, 1988). Individual learning is an important feature in static and homogeneous environments, because one individual can learn many things about the environment from a single interaction with it. However, if the environment is dynamic or heterogeneous, then that individual needs many interactions with the environment before it gets to know it. Because a single individual might not get enough chances to interact with such environment, cultural transmission (meaning learning from the experiences of others) becomes a requisite, too. In fact, individuals that have more chances to succeed in achieving their goals are the ones that combine both learning sources, thus increasing their gain in knowledge.

In order to solve any problem with PSO, we need to define a fitness function which will be used to measure the quality of possible solutions for that problem. Then, solving the original problem is equivalent to optimizing parameters of the fitness function, such that we find one of its minimum or maximum values (depending on the fitness function). By using a set of possible solutions, PSO will optimize our fitness function, thus solving our original problem. In PSO terms, each solution is called a particle and the set of particles is called a swarm. Particles gather and share information about the problem in order to increase their quality and hopefully become the optimum solution of the problem. Therefore, the driving force of PSO is the collective swarm intelligence (Clerc, 2006).

The fitness function generates a problem landscape in which each possible solution has a corresponding fitness value. We can imagine that similar to birds foraging in their environment, the PSO particles move in this landscape searching locations with higher rewards and exchanging information about these locations with their neighbours. Their common goal is to improve the quality of the swarm. During the search process the particles change their properties (location, speed, memory, etc) to adapt to their environment.

3. Particle Swarm Optimization and the Thompson problem

In 1904, while working on his plum pudding model of the atom, the British physicist Joseph John Thomson stated the following problem: what is the minimum energy configuration of $N$ electrons confined on the surface of a sphere? Obviously, each two electrons repel each other with a force given by Coulomb's law:

$$F = \frac{q_e^2}{4\pi\varepsilon_0} \frac{1}{d^2},$$

where $\varepsilon_0$ is the electric constant of vacuum, $q_e$ is the charge on a single electron, and $d$ is the distance between the two electrons. Because of these forces, each electron will try to get as far as possible from the others. However, being confined on the surface of the sphere, they will settle for a system configuration with minimum potential energy. The potential energy of a system with $N$ electrons is:
where we consider the electrons numbered in some random fashion, and $d_{ij}$ is the distance between electrons $i^{th}$ and $j^{th}$ in the system configuration $S$. Any configuration with minimum potential energy is called a “ground state” of the system.

Over the years, various generalizations for the Thomson problem have also been studied from different aspects. The most common generalization involves interactions between particles with arbitrary potentials. Bowick studied a system of particles constrained to a sphere and interacting by a $d^{-\gamma}$ potential, with $0 < \gamma < 2$ (Bowick et al, 2002). Travesset studied the interactions of the particles on topologies other than the 2-sphere (Travesset, 2005). Levin studied the interactions in a system with $N-1$ particles confined on the sphere and 1 particle fixed in the centre of the sphere (Levin & Arenzon, 2003). In general, finding the ground state of a system of $N$ repulsive point charges constrained to the surface of the 2-sphere is a long standing problem, which was ranked 7 in Stephen Smale’s famous list (Smale, 2000) of 18 unsolved mathematical problems to be solved in the 21st century, along with other famous problems, like the Navier-Stokes equations, Hilbert’s sixteenth problem and the P=NP problem.

Apart from physics, the Thomson problem arises in various forms in many other fields: biology (determining the arrangements of the protein subunits which comprise the shells of spherical viruses), telecommunications (designing satellite constellations, selecting locations for radio relays or access points for wireless communications), structural chemistry (finding regular arrangements for proteins S-layers), mathematics, economy, sociology, etc. From an optimization point of view, the Thomson problem is of great interest to computer scientists also, because it provides an excellent test bed for new optimization algorithms, due to the exponential growth of the number of minimum energy configurations and to their characteristics.

The Thomson problem can be solved exactly for small values of $N$ point charges on the surface of a sphere or a torus. However, for large values of $N > 8$, exact solutions are not known. The configurations found so far for such values display a great variety of geometrical structures. The best known solutions so far for such systems were identified with numerical simulations, using methods based on Monte Carlo simulations, evolutionary algorithms, simulated annealing, etc (Carlson et al., 2003; Morris et al., 1996; Perez-Garrido et al., 1996; Pang, 1997). PSO for the Thompson problem was first introduced in (Bătău & Bătău, 2007).

We will present in the following how the real-valued version of PSO can be used to tackle the Thomson problem. In order to avoid confusion, we will use the term “point charges” to refer to physical particles on the sphere (electrons, for example) and “particles” to refer to the data structures used by the PSO algorithm.

As mentioned in the previous section, in order to use a PSO algorithm we need to define a function that will measure the quality of solutions encoded by particles. One can think of many such functions for the Thompson problem, but a simple and quick solution is to use the potential energy of a system. We can save some computation time if we ignore the physical constants and use a simplified version of (2) for our fitness function:
where $N$ is the number of point charges in the system, $d_{ij}$ is the Euclidian distance between point charges $i$ and $j$, encoded by the particle $P$. If we represent our system configuration in 3D space using a Cartesian coordinate system, then we need to handle $3N$ real values, for the values on the Ox, Oy and Oz axis of each particle (see Figure 1). We will also need to explicitly enforce the sphere surface constraints which require additional computation time.

![Figure 1. Point charge represented in 3D Cartesian coordinate system](image1)

The memory requirements can be reduced and the computation overhead for constraint enforcing can be avoided, if we scale our system to the unit sphere and represent its configuration using a Spherical coordinate system. In this way, the sphere surface constraint is implicitly enforced and since $r$ is constant, the system configuration is encoded with only $2N$ real values, representing the azimuth $\phi$ and elevation $\theta$ angles (see Figure 2).

![Figure 2. Point charge represented in 3D Spherical coordinate system](image2)

In this case, the distance between point charges $i$ and $j$ located on the surface of the unit sphere is

$$d_{ij} = \sqrt{2 - 2[\cos \phi_i \cos \phi_j + \sin \phi_i \sin \phi_j \cos(\theta_i - \theta_j)]},$$  (4)
where $\phi_i, \phi_2 \in [-\pi, \pi]$ is the azimuth angle and $\theta_i, \theta_j \in [-\pi/2, \pi/2]$ is the elevation angle. Thus, PSO particles move in the search space $[0,1]^N$ and the location $x \in [0,1]^N$ of a particle decodes into a system configuration with:

$$\phi_i = 2\pi x_{2i-1} - \pi$$

$$\theta_i = \pi x_{2i} - \pi / 2$$  \hfill (5)

With this setup in place, the PSO algorithm begins with a swarm of particles randomly scattered around the search space. This generic initialization method could be replaced with a problem specific one (spherical initialization, for example). Each particle has a set of neighbours with which it will exchange information. An iterative process begins, which updates the properties of the particles. On each iteration each particle use the information from its own memory and the information gathered from its neighbours to update its properties. The equation used for updating the speed is:

$$v_t = \omega v_{t-1} + \phi_1 R_1 (p_{t-1} - x_{t-1}) + \phi_2 R_2 (g_{t-1} - x_{t-1}),$$ \hfill (7)

where $v_t$ is the speed at iteration $t$, $x_t$ is the location of the particle at iteration $t$, $p_t$ is the best location the particle has found until iteration $t$, $g_t$ is the best location the neighbours of the particle found up to the iteration $t$. The individual learning and cultural transmission factors ($\phi_1$ and $\phi_2$) control the importance of the personal and neighbour's experience on the search process. Note that although they share the same notation, these are parameters of the algorithm and are distinct and not related to the azimuth angles of the point charges. Because the importance of individual learning and cultural transmission is unknown, the learning factors are weighted by random values $R_1, R_2 \in [0,1]$. Usually the speed is bounded by some $v_{\text{max}}$ parameter to prevent it from increasing too much because of these random values.

With the updated speed vector and the old position of the particle, the new position is computed with:

$$x_t = x_{t-1} + v_t \Delta t,$$ \hfill (8)

for $\Delta t = 1$ iteration.

Based on the previous discussion, the PSO algorithm used for the Thomson problem is summarized in Figure 3. The algorithm is very simple and requires basic programming skills to be implemented in any programming language. It has many parameters that can be tuned in order to achieve high performance results. The tuning process of these parameters is beyond the purpose of this chapter. For now, let’s consider the following setup: $\omega = 0.9$ — will allow the algorithm to avoid rapid changes in the trajectories of the particles; $\phi_1 = \phi_2 = 2$ — gives equal weight to individual and social learning; iterations = 500 — for small and medium size systems, this should be enough for the particles to discover and focus on good solutions; $M = 2N$ — increases the swarm size with the size of the system.
Performing 10 runs of the algorithm from Figure 3 for systems with different sizes, we obtained the results presented in Table 1:

<table>
<thead>
<tr>
<th>N</th>
<th>Minimum known energy</th>
<th>Energy of PSO solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.500000000</td>
<td>0.500000000</td>
</tr>
<tr>
<td>3</td>
<td>1.732050808</td>
<td>1.732050808</td>
</tr>
<tr>
<td>4</td>
<td>3.674234614</td>
<td>3.674234614</td>
</tr>
<tr>
<td>5</td>
<td>6.474691495</td>
<td>6.474691495</td>
</tr>
<tr>
<td>6</td>
<td>9.985281374</td>
<td>9.985281374</td>
</tr>
<tr>
<td>7</td>
<td>14.452997414</td>
<td>14.452987365</td>
</tr>
<tr>
<td>8</td>
<td>19.675287861</td>
<td>19.675287861</td>
</tr>
<tr>
<td>9</td>
<td>25.759986531</td>
<td>25.759986599</td>
</tr>
<tr>
<td>10</td>
<td>32.716949460</td>
<td>32.717003657</td>
</tr>
<tr>
<td>15</td>
<td>80.670244114</td>
<td>80.685310397</td>
</tr>
<tr>
<td>20</td>
<td>150.881568334</td>
<td>150.953443814</td>
</tr>
<tr>
<td>25</td>
<td>243.812760299</td>
<td>243.898092955</td>
</tr>
<tr>
<td>30</td>
<td>359.603945904</td>
<td>359.901863399</td>
</tr>
<tr>
<td>35</td>
<td>498.569872491</td>
<td>499.018395878</td>
</tr>
<tr>
<td>40</td>
<td>660.675278835</td>
<td>661.666117852</td>
</tr>
<tr>
<td>45</td>
<td>846.188401061</td>
<td>847.129739052</td>
</tr>
<tr>
<td>50</td>
<td>1055.182314726</td>
<td>1056.517970873</td>
</tr>
</tbody>
</table>

Table 1. Minimum energies for Thomson problem found in experiments

From the results in Table 1, one can see that this simple PSO algorithm can provide high quality estimates for the ground states of various instances of the Thomson problem. The algorithm can be further improved not only in its parameters, but also in its structure (using a more advanced initialization method, for example). Obviously, the Particle Swarm Optimization algorithm can be applied for generalized forms of Thomson problem and other related problems, not only from Statistical Physics, but other domains, too.
4. Binary Particle Swarm Optimization and Ising Spin Glasses

Matter is composed of atoms and each atom carries a spin, meaning the magnetic moment of the microscopic magnetic field around the atom generated by the motion of the electrons around its nucleus.

If we heat a metal object higher than the Curie point of its material, the object will loose its ferromagnetic properties and become paramagnetic. At that point, the spins of the atoms change randomly so erratic that at any time they can point with equal probability to any possible direction. In this case, the individual microscopic magnetic fields generated by the spins cancel each other out, such that there is no macroscopic magnetic field (Huang, 1987). When the temperature is lower than the Curie point, in some metals (iron and nickel, for example) the spins of the atoms tend to be polarized in the same direction, producing a measurable macroscopic magnetic field. This is called “ferromagnetic” behaviour. By contrast, below the Curie point, in spin glasses only some pairs of neighbouring spins prefer to be aligned, while the others prefer to be anti-aligned, resulting two types of interactions between atoms: ferromagnetic and anti-ferromagnetic. Because of this mix of interactions, these systems are called disordered (den Hollander & Toninelli, 2005).

In the past, condensed matter physics has focused mainly on ordered systems, where symmetry and regularity lead to great mathematical simplification and clear physical insight. Over the last decades, spin glasses became a thriving area of research in condensed matter physics, in order to understand disordered systems. Spin glasses are the most complex kind of condensed state encountered so far in solid state physics (De Simone et al., 1995). Some examples of spin glasses are metals containing random magnetic impurities (called disordered magnetic alloys), such as gold with small fractions of iron added (AuFe).

Apart from their central role in Statistical Physics, where they are the subject of extensive theoretical, experimental and computational investigation, spin glasses also represent a challenging class of problems for testing optimization algorithms. The problem is interesting because of the properties of spin glass systems, such as symmetry or large number of plateaus (Pelikan & Goldberg, 2003).

From an optimization point of view, the main objective is to find the minimum energy for a given spin glass system (Hartmann, 2001; Pelikan & Goldberg, 2003; Fischer, 2004; Hartmann & Weigt, 2005). System configurations with the lowest energy are called ground states and thus the problem of minimizing the energy of spin glass instances can be formulated as the problem of finding ground states of these instances (Pelikan et al., 2006). The main difficulties when searching for ground states of spin glasses come from the many local optima in the energy landscape which are surrounded by high-energy neighbouring configurations (Pelikan & Goldberg, 2003).

The Ising model is a simplified description of ferromagnetism, yet it is extremely important because other systems can be mapped exactly or at least approximately to it. Its applications range from neural nets and protein folding to flocking birds, beating heart cells and more. It was named after the German physicist Ernst Ising who first discussed it in 1925, although it was suggested in 1920 by his Ph.D. advisor, Wilhelm Lenz. Ising used it as a mathematical model for phase transitions with the goal of explaining how long-range correlations are generated by local interactions.

The Ising model can be formulated for any dimension in graph-theoretic terms. Let us consider a spin glass system with $N$ spins and no external magnetic field. The interaction graph $G=(V,E)$ associated with the system has the vertex set $V = \{v_1, \ldots, v_N\}$. Each vertex
\( i \in V \) can be in one of two states \( S_i \in \{-1,1\} \). Edges in this graph represent bonds between adjacent atoms in the spin glass system. Each edge \( ij \in E \) has assigned a coupling constant, denoted by \( J_{ij} \in \{-J,J\} \); an edge exists between vertices \( i \) and \( j \) if the interaction between atoms \( i \) and \( j \) is not zero. In the classic model, this graph is a standard “square” lattice in one, two, or three dimensions. Therefore, each atom has two, four, or six nearest neighbours, respectively (see Figure 4). However, various papers present research done on larger dimensions (Hartmann, 2001).

**Figure 4.** Two dimensional Ising spin glass system

**Figure 5.** Ground state of the system from Figure 4 (values inside circles represent the states of the spins; dashed lines represent unsatisfied bonds)

For a system configuration \( S \), the interaction between neighbouring vertices \( i \) and \( j \) contributes an amount of \( -J_{ij}S_iS_j \) to the total energy of the system, expressed as the Hamiltonian:

\[
H(S) = - \sum_{ij \in E} J_{ij}S_iS_j .
\]  

The sign of \( J_{ij} \) gives the nature of the interaction between neighbours \( i \) and \( j \). If \( J_{ij} \) is positive, the interaction is ferromagnetic. Having the two neighbours in the same state \( (S_i = S_j) \) decreases the total energy. If \( J_{ij} \) is negative, the interaction between neighbours \( i \)
and \( j \) is anti-ferromagnetic. The decrease in total energy is obtained if they have opposite states. When all coupling constants are positive (or negative), a lowest-energy configuration is obtained when all vertices have the same state. This is the case of ferromagnetic materials. When the coupling constants are a mix of positive and negative values, as is the case for spin glasses, finding the “ground state” is a very difficult problem. A ground state of the system from Figure 4 is presented in Figure 5.

The two-dimensional Ising model of ferromagnetism has been solved exactly by Onsager (Onsager, 1944). The most common configurations in the literature are 2D Ising spin glasses on a grid with nearest neighbour interactions. In the case of no periodic boundary conditions and no exterior magnetic field, the problem reduces to finding a minimum weight cut in a planar graph for which polynomial time algorithms exist (Orlova & Dorfman, 1972; Goodman & Hedetniemi, 1973; Hadlock, 1975). Barahona showed that finding a ground state for the three-value coupling constant \( (J_{ij} \in \{-1,0,1\}) \) on a cubic grid is equivalent to finding a maximum set of independent edges in a graph for which each vertex has degree 3 (Barahona, 1982). He also showed that computing the minimum value of the Hamiltonian of a spin glass with an external magnetic field,

\[
H(S) = -\sum_{ij \in E} J_{ij} S_i S_j - h \sum_{i \in V} S_i S_i, \quad (10)
\]

is equivalent to solving the problem of finding the largest set of disconnected vertices in a planar, degree-3 graph. This means that finding ground states for three-dimensional spin glasses on the standard square lattice and for planar spin glasses with an external field are NP-complete problems. Istrail showed that the essential ingredient for the NP-completeness of the Ising model is the non planarity of the graph (Istrail, 2000).

Particle Swarm Optimization was introduced as a technique for numerical optimization and has proved to be very efficient on many real-valued optimization problems. Because finding the ground state of a spin glass system in the Ising model is a combinatorial problem, we need to apply a modified version of PSO. We will use the binary version of PSO (Kennedy & Eberhard, 1997). In this case, the \( i \)th component of the position vector of a particle encodes the state of the \( i \)th spin in the system (0 means down, 1 means up), while the \( i \)th component of the velocity vector determines the confidence of the particle that the \( i \)th spin should be up. On each iteration of the search process, each particle updates its velocity vector (meaning its confidence that the spins should be up) using (7). After that, the particle’s position vector (meaning its decision about spins being up or down) it updated using the component-wise formula:

\[
x_{ti} = \begin{cases} 
1, & \text{if } R < (1 + \exp(-v_{ti}))^{-1} \\
0, & \text{otherwise}
\end{cases} \quad (11)
\]

where \( R \in [0,1) \) is a random value. Once a particle’s position is known, its profit can be computed by:

\[
F(x_i) = -\sum_{1 \leq i < j \leq N} J_{ij} (2x_{ti} - 1)(2x_{tj} - 1), \quad (12)
\]
With such fitness function, lower values indicate better solutions. Obviously, this fitness function is inspired by the Hamiltonian given in (9) and can be adapted easily to external magnetic field environments using (10).

Based on the previous discussion, the Binary PSO algorithm used for the Ising spin glass problem is presented in Figure 6. A more advanced PSO algorithm for this problem is described in (Băutu et al., 2007). It combines the PSO algorithm with a local optimization technique which allows the resulting hybrid algorithm to fine tune candidate solutions.

1. Initialize $M$ random particles
2. for $t = 1$ to iterations
3. for each particle
   4. Update $v_t$ according to (7)
   5. Update $x_t$ according to (11)
   6. Evaluate $x_t$ using (12)
   7. Update $p_t$ and $g_t$ according to their definition
4. next
5. next
6. return solution from the particle with smaller fitness

Figure 6. PSO algorithm for the Ising spin glass problem

In order to test this algorithm, one can use a spin grass system generator, like the Spin Glass Server (SGS). SGS can be used to solve exactly 2D and 3D systems with small sizes or to generate systems for testing. It is available online at http://www.informatik.uni-koeln.de/ls_juenger/research/sgs/sgs.html.

<table>
<thead>
<tr>
<th>N</th>
<th>SGS minimum energy per spin</th>
<th>PSO minimum energy per spin</th>
</tr>
</thead>
<tbody>
<tr>
<td>64 / 3D</td>
<td>-1.6875</td>
<td>-1.6875</td>
</tr>
<tr>
<td>64 / 3D</td>
<td>-1.7500</td>
<td>-1.7500</td>
</tr>
<tr>
<td>64 / 3D</td>
<td>-1.8750</td>
<td>-1.8750</td>
</tr>
<tr>
<td>125 / 3D</td>
<td>-1.7040</td>
<td>-1.6720</td>
</tr>
<tr>
<td>125 / 3D</td>
<td>-1.7680</td>
<td>-1.7360</td>
</tr>
<tr>
<td>125 / 3D</td>
<td>-1.7360</td>
<td>-1.7040</td>
</tr>
</tbody>
</table>

Table 2. Minimum energies for Ising spin glasses found in experiments

Table 2 presents the energy per spin values obtained for 3D systems of 4x4x4 and 5x5x5 spins using (13). They will give you an idea about the performance of the binary PSO on this type of problems. The actual values depend on the spin system for which the algorithm is used.

$$E(x_t) = \frac{F(x_t)}{N}$$  \hspace{1cm} (13)

The results from table 2 were obtained without any tuning of the PSO parameters: the individual and social learning factors are $\phi_1 = \phi_2 = 2$ and the inertia factor is $\omega = 0.9$. The number of iterations is twice the number of spins, and the number of particles is three times
the number of spins. SGS provides the minimum energy for these systems using a branch-and-cut algorithm (De Simone et al., 1995).

5. Conclusions

This chapter presented the basic traits of Particle Swarm Optimization and its applications for some well-known problems in Statistical Physics. Recent research results presented in the literature for these problems prove that PSO can find high quality solutions in reasonable times (Băutu et al., 2007; Băutu & Băutu, 2008). However, many questions are still open: how do the parameters setups relate to the problems tackled? how can we improve the basic PSO to get state of the art results? how can we tackle very large size systems?

6. References


Particle swarm optimization (PSO) is a population based stochastic optimization technique influenced by the social behavior of bird flocking or fish schooling. PSO shares many similarities with evolutionary computation techniques such as Genetic Algorithms (GA). The system is initialized with a population of random solutions and searches for optima by updating generations. However, unlike GA, PSO has no evolution operators such as crossover and mutation. In PSO, the potential solutions, called particles, fly through the problem space by following the current optimum particles. This book represents the contributions of the top researchers in this field and will serve as a valuable tool for professionals in this interdisciplinary field.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: